Vision Res. Vol. 28, No. 6, pp. 695–709, 1988
Printed in Great Britain

0042-6989/88 $3.00 + 0.00

PERCEIVED CONTRAST AND STIMULUS SIZE: EXPERIMENT AND SIMULATION

MARK W. CANNON JR and STEVEN C. FULLENKAMP

H. G. Armstrong Aerospace Medical Research Laboratory, AAMRL/HEF Wright-Patterson Air Force Base, OH 45433 and Systems Research Labs Inc., 2800 Indian Ripple Road, Dayton, OH 45440, U.S.A.

(*Received* 13 *February* 1987; *in revised form* 15 *December* 1987)

Abstract—Perceived contrast functions were determined for three different Gabor patch sizes using magnitude estimation and verified by contrast matching. While thresholds show a significant decrease with decreasing patch size, perceived contrasts are equal and independent of patch size for contrasts above 0.06. Contrast matching was also used to study the apparent contrast of two other spatially limited stimuli; the sum of two orthogonal 4 c/deg sine waves multiplied by a gaussian envelope and the sum of spatially adjacent positive and negative gaussians. Models of contrast perception, based on tuned Gabor spatial filters, were formulated and tested for agreement with our experimental data. A model that pools filter responses across spatial frequencies and orientations was found to be more in agreement with our data than a model that simply uses the response of a single, maximally excited, mechanism to mediate contrast perception. Optimum filter bandwidth was found to be about 1.1 octaves.

Contrast perception; Spatial filtering; Visual modelling.

INTRODUCTION

While it is well known that threshold detection of sine wave gratings is strongly dependent on the number of cycles in the grating (e.g. Robson and Graham, 1981), two recent studies have demonstrated that the perceived contrast of suprathreshold gratings is relatively independent of the number of cycles. One study (Takahashi and Ejima, 1984) reported that apparent contrast remained constant as the stimulus size was reduced from 20 cycles to 1 cycle in either width or height for physical contrasts above 0.1. For contrasts below this level, apparent contrast begins to show a reduction for stimuli with small numbers of cycles until, for near threshold contrasts, the apparent contrast shows a steady decrease as the number of cycles is reduced from 10 to 1. The other study (Swanson *et al.*, 1984) performed at a contrast of 0.1 reported constancy of apparent contrast as the number of cycles was reduced from 16 to 2 but reported a decrease in apparent contrast as the width of the grating was reduced below 2 cycles in agreement with the low contrast Takahashi and Ejima data. If perceived contrast is mediated by a spatial array of mechanisms with antagonistic surrounds that multiply the input image as proposed by Wilson (1980) and by Swanson *et al.* (1984), the results mentioned above place restrictions on the spatial size of these mechanisms at high levels of physical contrast. These results imply that those mechanisms tuned to the grating spatial frequency must be only about one cycle wide at high contrasts since wider mechanisms tuned to the spatial frequency of the grating would show a decrease in their response (Kersten, 1984) and hence in perceived contrast as the stimulus approached 1 cycle in width. The Takahashi and Ejima study was apparently performed with abruptly truncated sine waves in both vertical and horizontal directions so their results are not free of contamination by higher spatial frequency harmonics. Swanson *et al.* used truncated cosine gratings which introduced low spatial frequency components as well (see Methods section). It is possible that these harmonics could have contributed to the constancy of contrast perception by exciting mechanisms tuned to spatial frequencies other than that of the grating. Since the dependence of apparent contrast on the number of stimulus cycles appears to have important implications for modelling suprathreshold contrast perception, we performed several experiments using both magnitude estimation and contrast matching to test this dependence. We used Gabor sine functions

for our stimuli in order to avoid the truncation problems mentioned above and to achieve optimal localization in both space and spatial frequency (Daugman, 1985).

We then attempted to quantitatively simulate these results using a two dimensional Gabor mechanism model for contrast perception. We found that in order to specify the free parameters in the model we had to run two more contrast matching experiments using two different stimuli; a sum of two spatially orthogonal 4 c/deg Gabor sines and a pair of small Gaussians, one positive and one negative with respect to the luminance of the background. Using the results of these experiments we determined what mechanism spatial size and what conditions of response pooling were required for a model of suprathreshold contrast perception to quantitatively match our suprathreshold experimental data.

METHODS

Apparatus

All experiments used computer generated imagery. Stimuli were generated as a 256 by 256 pixel array on a PDP-11 computer and were transferred to a Grinnell video frame buffer capable of storing three images for display. Contrast of the video images was determined by an analog multiplier controlled by an AIM 65 microcomputer. The contrast controlled video signal was fed to our Tektronix 632 (P4 white phosphor) monitor after final modification by a circuit to assure linearity of the monitor luminance as a function of input voltage. Monitor luminance linearity was confirmed and the system calibrated with a Pritchard Spectra photometer. All experiments were performed at an average luminance of 170 cd/m^2 with a stimulus presentation time of 500 msec. Stimulus rise and fall times were abrupt. The screen was masked off with a large white cardboard surround to an area 6.5 deg square. The surround was illuminated by white light, matched to the luminance of the screen, over an area 24 deg high by 31 deg wide.

Stimuli

The luminance profile of the Gabor functions used in these experiments is given by

$$L = L_0 + A(\exp - [(x/T)^2 + (y/T)^2]/\sigma_s^2)\sin(2\pi x/T) \quad (1)$$

where x and y are distances from the origin in degrees of visual angle, T is the period of the sine wave in degrees and L_0 is the average luminance. The spatial frequency subtended at the subject viewing distance was 4 c/deg for all patterns. The term σ_s denotes the number of stimulus cycles between the center of the function and the point where the guassian envelope reaches $1/e$. The use of a Gabor sine function assured that the average luminance of the pattern would remain constant, independent of the width of the gaussian envelope. It is probably not widely appreciated that the spectrum of a Gabor cosine function accumulates a steadily increasing d.c. component as the width of the function σ_s decreases. This can be visualized rather easily by considering what happens when σ is reduced to a small fraction of a period. Under these conditions the negative portions of the cosine beyond a quarter period are attenuated to near zero while a significant region of the positive portion near the origin is preserved at near the original amplitude. Since the Gabor sine function is anti-symmetric across the origin, no d.c. component arises no matter how narrow the envelope becomes.

As the diameter of the Gabor sine function decreases, the maximum peak to trough amplitude of the sine wave, under the gaussian envelope, also decreases. If we specified our contrast by A/L_0 in equation (1), the actual peak to trough contrast seen by the observer would decrease with decreasing sigma. Therefore, we calibrated our system so that the contrast controlled by the computer was the usual definition of contrast, $(L_{max} - L_{min})/(L_{max} + L_{min})$, where L_{max} and L_{min} were the measured luminances at the highest peak and lowest trough of whatever Gabor pattern was presented.

Magnitude estimation

The three stimuli used in the magnitude estimation experiment are shown in Fig. 1. Each of these stimuli were presented, in random order, at 12 different, logarithmically spaced, contrast levels in a single experimental session. Contrasts ranged from 0.005 to 0.5 and each contrast level was presented twice in the session. The seven subjects, who had all performed magnitude estimation experiments before, were told to call out numbers proportional to the apparent contrast at the center of each grating as it appeared (Cannon, 1984, 1985). These numbers were recorded by the experimenter. This is the free modulus method of magnitude estimation

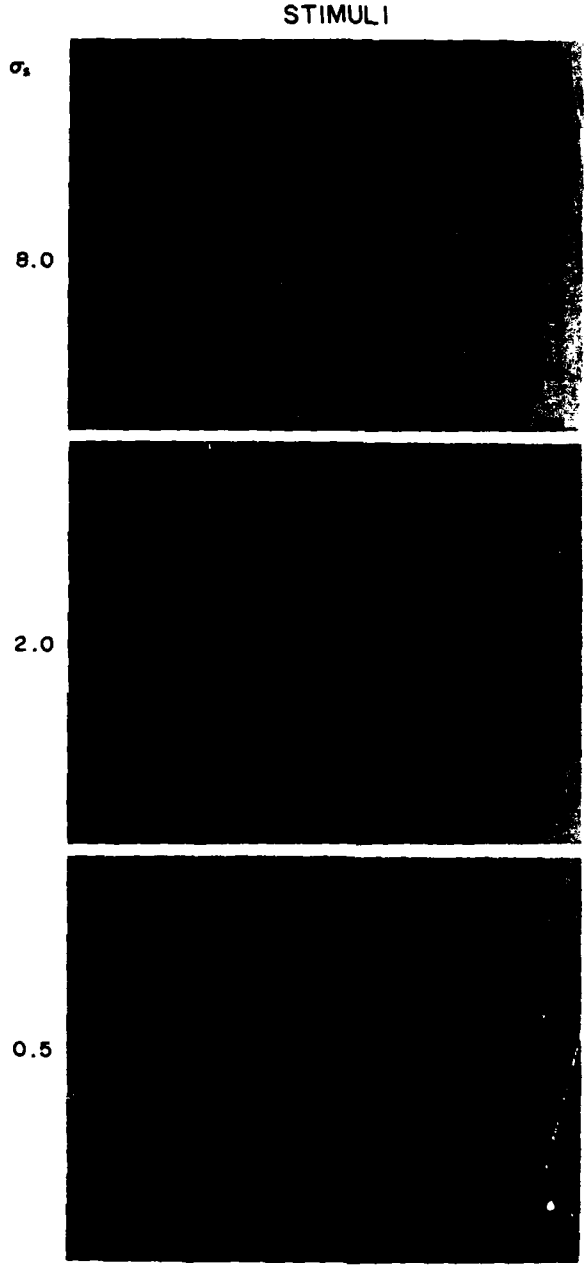

Fig. 1. Gabor stimuli used in the magnitude estimation experiments. Note that in the stimulus for $\sigma = 0.5$ only about 1 cycle of the sine wave is visible.

(Stevens, 1956) in which no standard is shown and no number scale is suggested to the subjects. During the 6 sec that elapsed between presentations the screen remained at the average luminance of 170 cd/m^2.

When the free modulus method of magnitude estimation is used, different subjects may use different number scales so some normalization is required to bring all estimates into a common range. This is usually accomplished by adjusting the geometric mean of the perceived contrast estimates from each subject to the same value. The adjustment is accomplished by determining the grand geometric mean for all perceived contrast estimates from all subjects and then determining a correction factor for each subject that will adjust the geometric mean of the individual data set up or down until it is equal to the grand mean. A mean perceived contrast is then determined for each level of physical contrast by computing the geometric mean, across subjects, of all the adjusted perceived contrast estimates at that level of physical contrast. Finally, a plot of these means vs physical contrasts on log–log coordinates defines a perceived contrast function. In order to also maintain the correct relative amplitude relationships among the perceived contrast estimates for stimuli with three different σ we only performed the normalization and determined the individual subject correction factors for the $\sigma_s = 8.0$ data. The same correction factors were then used to normalize that subject's perceived contrast estimates obtained for $\sigma_s = 2.0$ and 0.5.

Contrast matching

Contrast matching experiments were performed to verify the magnitude estimation results at a contrast of 0.3. In these experiments, the apparent contrast of the smallest grating patch ($\sigma_s = 0.5$) was compared with the perceived contrast of a vertical full screen grating. Both of these vertically oriented 4 c/deg gratings were matched in contrast to a 5.5 c/deg horizontal grating. The experiments were performed in this way to eliminate the possibility of adaptation effects that can occur while matching two gratings of the same spatial frequency and orientation. The method used was a variation of the method of constant stimuli. One of the gratings was set to a contrast of 0.3. This was the comparison grating. The other grating (the test grating) was presented at 8 different contrast levels spaced 1 dB apart and bracketing the contrast of subjective equality. Stimuli were displayed for 500 msec with abrupt rise and fall times as in the magnitude estimation experiments. In a single trial, the test and comparison gratings were presented in sequential 1 sec intervals marked by auditory tones and subjects indicated, by means of a switch position, which interval contained the stimulus of higher contrast. The order of presentation varied randomly from trial to trial. A complete set of data from each subject consisted of responses to 30 trials at each of the 8 contrast levels. These data could then be plotted in the form of psychometric functions showing the probability that the test grating has higher apparent contrast than the comparison grating vs the contrast of the test grating. The contrast of subjective equality is the contrast at which the psychometric function crosses the 0.5 probability level.

Threshold measurements

All thresholds were measured using a temporal two-alternative forced-choice procedure. Stimulus durations were the same as those used in the suprathreshold portions of the experiment and stimulus intervals were marked by auditory tones. Contrasts were decreased by 2 dB after three correct responses and increased by 2dB after 1 incorrect response. The staircase was stopped after 10 contrast reversals. Threshold was the log mean of the 10 contrasts at which reversals occurred. Mean thresholds for each patch size were computed by taking the geometric mean of all individual thresholds.

RESULTS

Magnitude estimation

Perceived contrast functions for the three different Gabor functions are shown in Fig. 2. The data points are geometric means of the normalized perceived contrast estimates from seven subjects. Since each subject made contrast estimates for all three Gabor functions in the same experimental session, we assume that all estimates were made using the same perceptual scale for contrast. The vertical arrows at the foot of each perceived contrast function are the mean thresholds for the seven subjects. The error bars on the data symbols beside the legend indicate the average size of the standard deviation, across subjects, for each patch size. It can be seen that subject differences are somewhat greater for the smaller patch. As expected, thresholds decrease as patch size increases but

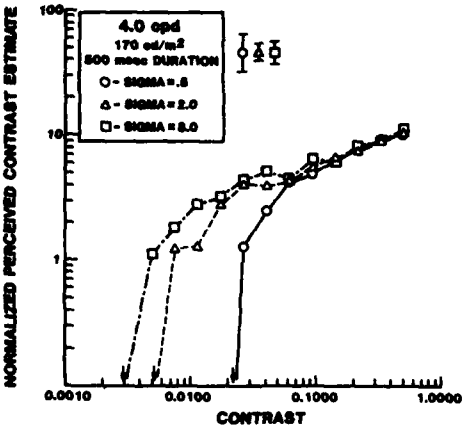

Fig. 2. Mean perceived contrast functions for seven subjects from free modulus magnitude estimation. Error bars next to the legend, at the top of the figure, are inter-subject standard deviation averaged across all contrast levels. The thresholds decrease with decreasing σ, as expected, but the perceived contrasts are equal for all stimuli above a contrast of about 0.1.

above a contrast of 0.06 perceived contrasts for all three patch sizes are equal. Our magnitude estimation results indicate that spatial summation has no effect at suprathreshold contrast levels. Perceived contrast perception, at least for sine wave gratings, is a spatially localized process.

Contrast matching

The contrast matching results are illustrated in Fig. 3. These experiments were performed to determine if the equality in perceived contrast for physical contrasts above 0.06 would be confirmed by a more classical psychophysical method. The matches from which the data in each panel are derived are indicated by the symbols for patch and full screen gratings shown at the top of each of the large panels. The two large panels in the left half of the figure are psychometric functions for four subjects obtained when they matched the vertical $\sigma_x = 0.5$ Gabor patch and a vertical full screen grating, both at 4 c/deg, to a horizontal full screen 5.5 c/deg comparison grating which had a fixed contrast of 0.3. The large right hand panels show the results of the reverse procedure where the patch and vertical full screen grating were fixed at a contrast of 0.3 and the horizontal 5.5 c/deg grating was the test grating. The vertical dashed line in each panel marks the physical contrast of the comparison grating. The horizontal dashed line drawn at a probability of 0.5 marks the level at which half the test stimuli are seen as greater than and half are seen as less than the comparison stimulus in apparent contrast. The test contrast at which the psychometric function crosses the horizontal dotted line is, therefore, the contrast of subjective equality.

A ratio of the contrast of the horizontal grating to the contrast of the vertical grating, at subjective equality, was calculated for each of the 8 psychometric functions in the two upper large panels, regardless of which was the test and which was the comparison grating. These ratios were then expressed in log units and a mean ratio was determined by averaging the individual log unit values and taking the antilog of the result. This ratio characterized the contrast ratio required for a match between the vertical and horizontal full screen gratings. The mean contrast ratio and its standard deviation are shown in the small upper right-hand panel in Fig. 3. The mean ratio was 1.005, demonstrating that the perceived contrasts of the vertical and horizontal full screen gratings were equal for equal physical contrasts. A similar calculation was made for the small patch vs the horizontal full screen matching data using the 8 psychometric functions in the two large lower panels. The results are shown in the small lower right hand panel. The ratio was 1.002, demonstrating that the perceived contrast of the patch and horizontal gratings were equal for equal physical contrasts. These equalities imply that there is no difference in perceived contrast between the small grating patch and the vertical full screen grating when both have the same physical contrast, at least for physical contrasts near 0.3. Overall, the matching results agree with our magnitude estimation results for high suprathreshold contrasts and confirm our assertion that perceived contrast is a spatially localized process.

DISCUSSION OF EXPERIMENTAL RESULTS

Our experiments, performed with Gabor functions, confirm the fact that the high contrast matching results of Takahashi and Ejima (1984), for small numbers of stimulus cycles, were not an artifact of the abrupt truncation of their stimuli. Our magnitude estimation data also show that equality in perceived contrast occurs only above a physical contrast of about 0.06, in agreement with their matching results. Careful study of our magnitude estimation data

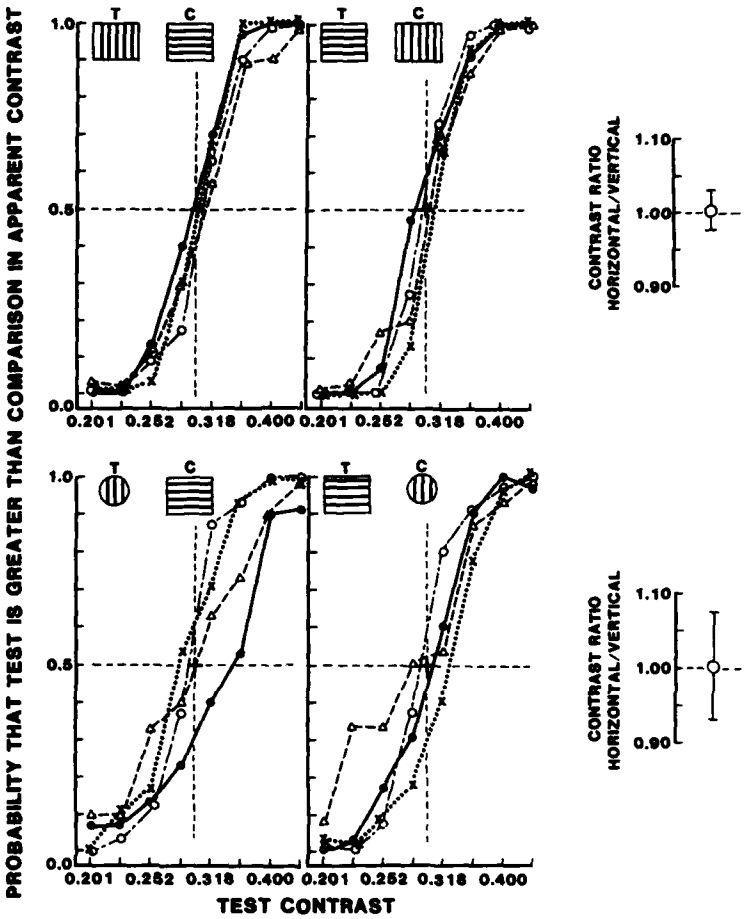

Fig. 3. Psychometric functions for four subjects from contrast matching experiments where a small patch and vertical full screen grating were matched in apparent contrast to a horizontal full screen grating using the method of constant stimuli. Each point represents 30 responses. The symbols above each panel show which stimulus was the variable contrast test and which was the comparison grating fixed at a contrast of 0.3. The small panels on the right show the ratio of contrasts for which the horizontal and vertical gratings attained subjective equality. See text for details.

(Fig. 2) reveal that the perceived contrast functions rise rather smoothly from threshold and merge at a contrast of about 0.06. Above this level all perceived contrast functions behave as power functions of contrast with an exponent of about 0.4. One of us (Cannon 1979, 1984, 1985) has repeatedly demonstrated that perceived contrast functions can be described as power functions of contrast minus threshold and the functions shown here are no exception. The exponents and multiplicative constants derived from least square fits to the data plotted in coordinates of log perceived contrast vs log contrast minus threshold are given in Table 1.

All three data sets are well fitted by similar functions. The estimates of the exponents vary only by a factor of 1.2 while the estimates of the multiplicative constants vary only by a factor of 1.07. While this type of fit is useful for descriptive purposes, it does not provide a description of the mechanisms that may be involved in contrast perception. After analyzing magnitude estimation data for brightness of circular disk targets, Drum (1986) has recently proposed that perceived brightness functions, while they can be fitted with threshold corrected power func-

Table 1. Perceived contrast function parameters

σ_m	Slope	Multiplicative constant
0.5	0.452	15.553
2.0	0.411	14.515
8.0	0.379	14.653

tions of the type described above for contrast, may also be described as the sum of responses from two different mechanisms. The response of one of these mechanisms rises fairly steeply with luminance and saturates about 0.5–1.5 log units above threshold. The second mechanism response rises much more gradually and although it shows a compressive non-linearity it continues to rise for luminance 2.5–3 log units above threshold.

A two mechanism model for contrast perception would also be consistent with the data presented in Fig. 2 of the present paper. The data in Fig. 2 show a very tight clustering in perceived contrast for physical contrasts above 0.06, confirmed by the matching experiment, so that all responses in this contrast range can be described by a single function that is independent of stimulus size. Below a contrast of 0.06, both threshold and the shape of the lower portion of the perceived contrast function show a strong dependence on the radius of the stimulus grating patch. This behaviour is consistent with the presence of a second mechanism with a size dependent threshold and a response that saturates at a perceived contrast level of about 5.

Although both our contrast data and Drum's brightness data can be modeled using two mechanisms with different ranges in the sensory domain, there are specific differences between contrast and brightness response. One of the most dramatic differences is that, unlike our perceived contrast data, Drum's brightness data show a change in the steepness of the upper branch of the curve with spot diameter, so responses of both of his mechanisms are influenced by stimulus size.

The nature of the mechanism that can produce the type of low contrast behavior observed here is not clear. Normally the difference in detection threshold due to stimulus extent are explained by probability summation (e.g. Robson and Graham, 1981) and this explanation has been adequate as long as only detection was considered. In Fig. 2, however, we see that, even when the grating patch is quite visible, its apparent contrast is less than that of a grating with more cycles as long as its physical contrast is less than 0.06. The matching data of Takahashi and Ejima (1984) show a similar effect. Clearly, for physical contrasts below 0.06, the actual response of mechanisms mediating contrast perception are lower when the number of cycles is lower. This implies mechanisms that actually compute a weighted sum of inputs from a large, but well defined, spatial region. Probability summation models, on the other hand, address only the initial detection of stimulus. They say nothing about stimulus appearance once threshold has been exceeded. It is likely that many spatial summation effects attributed to probability summation can be also modelled by a hardwired spatial summation model with judicious adjustment of the summation weights. However, it is also likely that detection depends on both. Our reason for proposing this duality is the extremely large range of 64 stimulus cycles over which spatial summation was found to act by Robson and Graham (1981). Based on current evidence, it appears unlikely that a single cortical mechanism would have a receptive field of this size. Thus, understanding the perceived contrast response in the region between threshold and a contrast of 0.06 will require more research. However, the behavior of mechanisms mediating contrast perception on the upper portion of the perceived contrast curve appears to be more straightforward. In this region, perceived contrast can be attributed to one type of mechanism with a small spatial summation area. As far as we know, no previous attempt has been made to simulate suprathreshold contrast perception with a two dimensional model. Consequently, we felt that some valuable insights could be obtained by attempting to develop such a model using spatial filter mechanisms of a type already described in the literature. We will show in the next section of this paper that estimates of the size and spatial frequency bandwidth of this mechanism can be inferred from a model of contrast perception incorporating mechanisms tuned to both spatial frequency and orientation.

A MODEL FOR CONTRAST PERCEPTION

In recent years, researchers have attempted to give mathematical descriptions for both cortical receptive fields and mechanisms that can explain psychophysical data relating to spatial pattern analysis in the human visual system. Several mathematical functions have been proposed, and they all show broad similarities in shape, consisting of both excitatory and inhibitory regions that are reduced in strength with distance from the center of the mechanism. In the neurophysiological literature, one finds attempts to model receptive field profiles divided, primarily, between difference of gaussian models (e.g. Parker and Hawken, 1985) and Gabor

function models (e.g. Webster and De Valois, 1985) while Young (1985) has proposed a Gaussian derivative model to explain certain data. The functions most prominent in the psychophysical literature are also difference of gaussians (e.g. Wilson and Bergen, 1979; Swanson *et al.*, 1984) and Gabor functions (e.g. Watson, 1982; Daugman, 1985), but others such as the Cauchy function (Klein and Levi, 1985) have also been proposed.

The difference of Gaussian model has been used to develop a detailed one dimensional model for both contrast detection (Wilson and Bergen, 1979) and some aspects of suprathreshold vision (Swanson *et al.*, 1984). The Gabor function was used by Watson (1982) to formulate a detailed one dimensional model for contrast detection that was later extended to a two dimensional threshold model (Watson, 1983).

As far as we can determine, there is no clear evidence that difference of gaussian mechanisms show more predictive power than Gabor mechanisms in any of the threshold models mentioned above. Consequently, we chose Gabor functions to represent the spatial filter mechanisms in our model because of their greater mathematical tractability for the type of two dimensional model we envisioned. They are describable by only two parameters, the half-width of their gaussian envelope and their spatial frequency. Furthermore, a closed form mathematical equation for their bandwidth has been derived (Watson, 1982).

We have chosen to ignore radially symmetric (mexican hat) filters because the preponderance of neurophysiological evidence for cortical filters implies orientation selectivity.

Block diagrams describing the spatial filtering stages of the models studied in this paper are similar to other models of visual pattern processing in the literature (Wilson and Bergen, 1979; Watson, 1982; Swanson *et al.*, 1984). These models assume that the visual system contains an array of filters, each tuned to a particular spatial frequency range and each operating on the input pattern over a particular spatial region of the image. The models we propose are shown in Fig. 4 and were simulated on a Definicon Systems 32 bit 68020 high speed co-processor board resident in an IBM PC-XT computer. Both models used identical filter and nonlinearity stages and these are shown in the left most panel of Fig. 4.

All tuned filters in the model were represented in the space domain by two dimensional Gabor sine functions (Watson, 1982) with center spatial frequencies 1/2 octave apart. All filter functions were energy normalized in the space domain so the two dimensional integral of the squared function over space equalled one. This normalization assures that the gain of all the filters will be equal at their center spatial frequency. Equality in gain across spatial frequencies guarantees model responses in good agreement with the relatively flat equal perceived contrast contours found at suprathreshold levels by Georgeson and Sullivan (1975) and Cannon (1985), at least between 2 and 8 c/deg. The lowest filter center frequency used in this study was 1 c/deg because responses of filters with lower center frequencies were negligible for all the stimuli we used. No restrictions on the existence of mechanisms tuned to higher or lower spatial frequencies than those shown are implied. The multiple paths emerging from each spatial frequency filter represent the fact that the image was filtered by multiple representations of the basic Gabor function with orientations that ranged from 0 to 165 deg in 15 deg steps. Filter responses from 180 to 345 deg were mirror images of those already computed.

The spatial representation of each filter was convolved with the input image (using fast fourier transform techniques) over a 4×4 deg area resolved into 256×256 pixels. This operation generated a different 256×256 pixel filtered image as the output of each filter. We have assumed uniform and equal density for all the mechanisms over the 4×4 deg area so the amplitude of each pixel of these filtered images represented the output of a spatial filter mechanism centered at that pixel location. The absolute value of this output was then subject to a non-linear power function transformation. The exponent of the power function was 0.5 since this is close to the exponent for the upper branch of our perceived contrast function shown in Fig. 2 and in good agreement with exponent values determined in other studies (Gottesman *et al.*, 1981; Cannon, 1984, 1985). The 256×256 pixel array of nonlinear responses produced by each filtering operation represented the spatial contrast information reported to some central processing center by the spatial array of mechanism composing the filter. In Fig. 4, these filter response arrays are labelled $F_{f,\theta}(x, y)$.

We simulated two possible computational

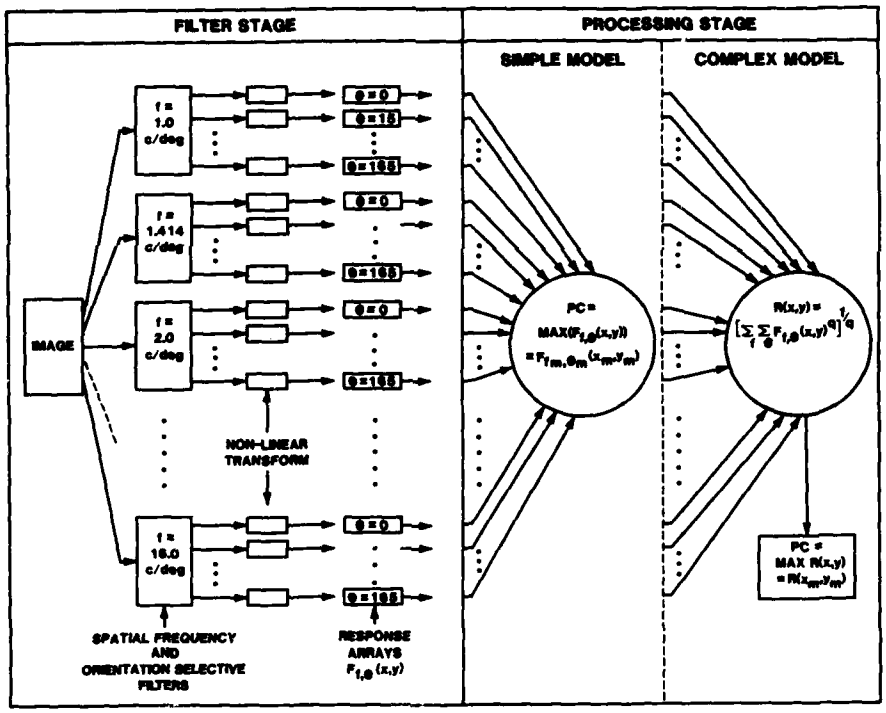

Fig. 4. Two possible models for contrast perceptioa. Both models share the same bank of orientation selective spatial filters. Each path emerging from a spatial frequency filter represents the response of a Gabor filter rotated to one of the allowed orientations between 0 and 165 deg. Orientation increments are 15 deg. Different summation rules for the two models are shown in the large circles. PC is the response that mediates the sensation of perceived contrast in the model. Details explained in text.

schemes for the extraction of perceived contrast information by the central processor. Since we have ruled out spatial summation across mechanisms as an important contributor to perceived contrast, one logical scheme would be that shown in the middle panel of Fig. 4. Here, the central processor scans all the filter response arrays and chooses the largest single pixel amplitude (the response of the single most strongly excited mechanism) to mediate the sensation of contrast. This will be referred to as our simple model.

A second model (the complex model) is represented in the right-hand panel of Fig. 4. In this model, the central processor pools responses from all filter response arrays using the Quick summation rule (Quick, 1974). Note that the response arrays are summed across spatial frequency and orientation at each spatial location (x, y) and that this summation is independent of all other spatial locations. Thus, no spatial summation across the x, y coordinates occurs in this model either. The pooling produces an intermediate response array, $R(x, y)$, Each element of $R(x, y)$ contains the summation of activity from all spatial frequency and orientation mechanisms centered at each location x, y. The central processor then scans the responses in the R-array and chooses the array member with the highest amplitude to mediate the sensation of contrast. In Fig. 4, the location of this point is denoted by X_m, Y_m.

There are two free parameters for the complex model. These are the half width of the Gabor filter mechanism gaussian envelope (σ_m) and the exponent, n, of the Quick summation, Q. The only free parameter for the simple model is σ_m. The object of simulations reported in this paper was to determine for what ranges of these parameters, if any, our models could predict experimental results. Our first modelling attempt was a comparison of the perceived contrast of a small stimulus patch ($\sigma_s = 0.5$) and a full screen grating; the same stimuli we addressed in the Results section of this paper. We chose σ_m values ranging from 1.4 to 0.14. The

values of the Quick exponent used in the simulation ranged from 2 to 4. Model responses are shown in Fig. 5.

An amplitude of 1.0 represents the model response to a full screen grating with a contrast of 0.3. The solid horizontal line represents the model response when the full screen grating is increased slightly in contrast to match the relative contrast of the full screen grating (1.002) required to match the contrast of the small ($\sigma_s = 0.5$) grating patch in our matching experiments. This horizontal line will be referred to subsequently as the expected response. The dotted horizontal lines represent model responses when the full screen grating contrast was increased or decreased to the levels defined by the standard deviation shown in Fig. 3. These lines are provided to give some sense of when the model would generate perceived contrast estimates that would be detectably different from experimental predictions. The points and connecting curves represent complex model responses to the small patch for the various values of the Quick exponent, Q. The asterisks represent the relative responses of the simple model.

Before discussing the results displayed in Fig. 5 there are two points regarding the model responses that should be addressed. First, some readers may question how the horizontal lines representing the relative model responses from our matching data can remain the same for all values of the Quick parameter displayed in the figure. For those readers, we show in Appendix 1 that this relative response is not affected by the Quick parameter, Q, but by the power function exponent of the nonlinear transformation. Second, if the reader is wondering why the response is not always maximized for a model receptive field matched in shape to the stimulus ($\sigma_m = 0.5$) let us assure you that no principles of matched filtering are being violated. Optimal matched filtering requires that all input stimuli be energy normalized and that is not the case here. Instead, all input stimuli have the same peak to trough amplitude implying equal contrast.

If we assume that the space between the dotted lines in Fig. 5 defines a region where the model response to a small patch is indistinguishable from the model response to a full screen grating, this figure clearly eliminates a wide range of parameters from contention. Note that the simple model response only approaches this range for σ_m of 0.28 or less. This corresponds to a spatial frequency bandwidth at half amplitude of 4.95 octaves or greater, a result at variance with current findings on cortical mechanism bandwidths (De Valois *et al.*, 1982). However, there is still a surprisingly large range of σ_m and Q that allows the complex model to produce valid responses for this stimulus. These are represented by the range of coordinates where the functions represented by the lines connecting the simulation points run between the dotted lines. It should be clear that there is a continuum of these parameter pairs. In an attempt to narrow down our choice of model parameters we performed contrast matching experiments and simulated model responses with two other stimuli.

Further matching experiments

The contrast matching and data analysis in these experiments used exactly the same techniques described in the Methods section of this paper. One stimulus was a pair of Gaussians, one positive and one negative with respect to the average luminance. This will be referred to

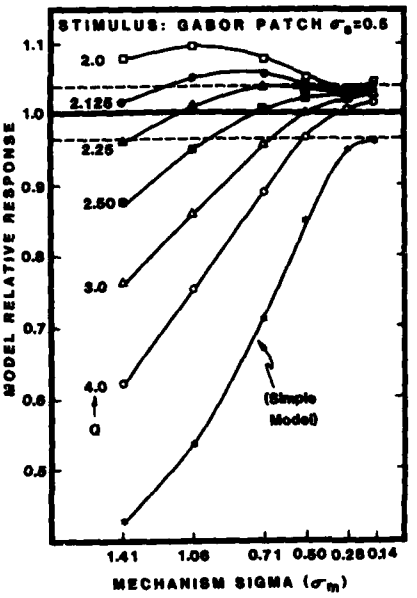

Fig. 5. Parametric study of model responses to a small $\sigma_s = 0.5$ Gabor patch. A response amplitude of 1.0 represents the model response to a full screen grating. The horizontal line just above 1.0 represents the expected model response to the small Gabor patch determined from our contrast matching experiments. The circles, squares and triangles represent complex model responses to the small Gabor patch for various values of Q, the Quick summation constant, and mechanism width, σ_m. Simple model responses are represented by the asterisks. Note that many combinations of these two parameters generate responses that agree with the expected response.

subsequently as the bi-polar gaussian stimulus. The Gaussians were spaced at a peak to trough distance equivalent to half of a 4 c/deg period. The sigma parameter for each Gaussian was one quarter of a 4 c/deg period. This stimulus was matched in contrast to a full screen sine wave grating by the same 4 subjects who participated in our earlier matching experiments.

The stimulus in the second experiment was a sum of two spatially orthogonal 4 c/deg sine wave gratings multiplied by a Gaussian envelope with a sigma parameter of 1.0. This grating was matched in contrast to a 4 c/deg Gabor rotated to an angle of 45 deg. The sigma parameter of the rotated grating patch was also 1.0. Three of the four subjects in this experiment were the same as those who participated in the earlier matching experiments.

Both of these stimuli were found to have slightly lower apparent contrast than the sine wave function they were matched to.

Further simulations

The bi-polar Gaussian simulation results and the orthogonal sine wave simulation results are shown in Fig. 6a and b respectively. The format of these panels is the same as the format of Fig. 5. The horizontal solid line in Fig. 6a represents the model response when a full screen grating is lowered in contrast by the amount required to match the bi-polar Gaussian. The horizontal solid line in Fig. 6b represents the model response when a $\sigma_s = 1.0$ Gabor patch is lowered in contrast by the amount required to match the orthogonal sine wave patch. The dotted lines represent model responses with contrasts adjusted up or down to match the standard deviations in contrast for each experiment.

The simple model would require mechanisms smaller than those considered in the figure and hence, inordinately wide bandwidths to generate responses that come sufficiently close to the expected response.

Finding the best parameters

The shapes of the functions generated by connecting the model response points for the complex model are different in the two panels of Fig. 6 and are different from those illustrated in Fig. 5. Some functions that intersected or ran close to the expected response line in Figs 5 and 6a do not do so in Fig. 6b. Some functions intersect or approach the expected response line at different values of σ_m for all three simulations. Studying the pattern of these functions revealed that a Q value of 2.5 and a σ_m value of 0.71

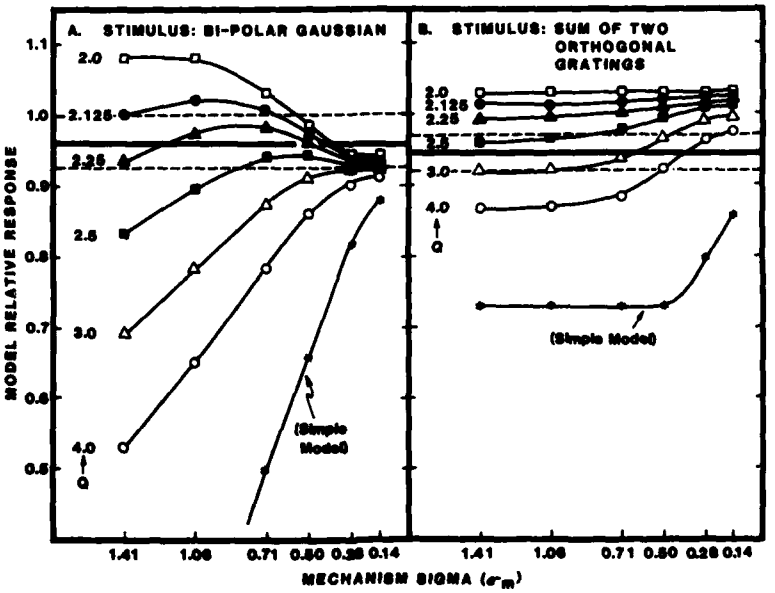

Fig. 6. Model responses to two different stimuli. The left-hand panel shows model responses for a bi-polar gaussian stimulus while the right-hand panel shows model responses to the sum of two orthogonal sine wave gratings multiplied by a gaussian envelope. The responses to the bi-polar gaussian are relative to the model response to a full screen grating. The responses to the orthogonal sine waves are relative to the model response to a Gabor sine patch with $\sigma_s = 1.0$. See text for details.

produced responses near the expected response line for all three patterns. This implies that there may be some Q, σ_m pair for which the distance to the expected response line is a minimum when averaged for all three simulations. Since we do not have an analytic form for these functions there is no way to exactly determine the specific parameter pair. Instead, we have chosen a method which should give us a good approximation to their true values. We computed the distance in log units between each model response point in Figs 5 and 6 and its respective expected response line. We then computed an RMS value of the log unit differences across the three simulations for each Q, σ_m pair indicated by the data points in Figs 5 and 6. These RMS values are summarized in Table 2. The table shows a clear minimum (outlined by a box) for a Q of 2.5 and a σ_m of 0.71 among the complex model results. The table also makes it clear that the simple model must be rejected because of its large RMS difference from the expected response. Thus, the complex model with the parameters stated above is an excellent predictor of perceived contrast for all the stimuli considered in this paper. The maximum error in predicted perceived contrast, using these parameters, is 0.016 log units.

The spatial frequency bandwidth, at half amplitude, of the filters used in this study allow comparisons with bandwidth estimates from quantitative models of threshold vision. The half amplitude bandwidths of our filters when $\sigma_m = 0.71$ is 1.13 octaves. The bandwidth proposed by Watson (1982, 1983) for the "sensors" in his threshold models of spatial pattern processing was about 0.5 octaves. The threshold model proposed by Wilson and Bergen (1979), on the other hand, used mechanisms with a bandwidth of about 1.75 octaves or higher. Our bandwidths fall almost midway between these two estimates. Thus we cannot, at present, make any inferences about differences in bandwidth between threshold and suprathreshold mechanisms.

Bandwidths greater than an octave have appeared in at least three studies of suprathreshold vision. In masking experiments, Legge and Foley (1980) found bandwidths of about 1.75 octaves near 2.0 c/deg while Wilson *et al.* (1983) found bandwidths that varied with spatial frequency from 2.5 octaves near 1.0 c/deg to 1.25 octaves near 16.0 c/deg. Experiments to measure the characteristics of spatial summation in suprathreshold noise by Kersten (1984) led to a bandwidth estimate, at half amplitude, of about 1.7 octaves at 8 c/deg. All of these bandwidth estimates are higher than ours but they were inferred directly from detection thresholds in experiments requiring the detection of a one dimensional stimulus against a masking background. Our bandwidths, on the other hand, are inferred from two dimensional model responses required to simulate subject judgements about the apparent contrast of several two dimensional stimuli. It would not be surprising if somewhat different bandwidth estimates resulted. It is also possible that the masking experiments and our contrast perception experiments investigated different mechanisms. In fact, Kersten's (1984) data appear to support this idea. Our magnitude estimation data show that perceived contrast is essentially independent of patch size for contrasts above 0.06 at 4 c/deg. Kersten's data for 2 and 8 c/deg (subject D.K.) show that detection threshold of a Gabor cosine function in noise is reduced by increasing its width. The stimulus contrasts in both cases are in a range where perceived contrast would not change with stimulus size (0.08–0.3) implying that contrast perception of a noise free signal and signal detection in a masking environment may be mediated by different mechanisms.

Swanson *et al.* (1984) found that "response pooling" across spatial frequency selective

Table 2. RMS difference between expected and achieved model response for various combinations of σ_m and Q

Q	σ_m 1.41	1.06	0.71	0.50	0.28	0.14
2.0	0.0408	0.0428	0.0341	0.0262	0.0244	0.0262
2.125	0.0200	0.0261	0.0253	0.0226	0.0228	0.0229
2.25	0.0180	0.0142	0.0179	0.0193	0.0223	0.0228
2.50	0.0501	0.0233	**0.0111**	0.0154	0.0204	0.0214
3.0	0.1067	0.0637	0.0269	0.0154	0.0178	0.0183
4.0	0.1920	0.1231	0.0614	0.0289	0.0178	0.0163
Simple model	0.3925	0.3083	0.1967	0.1223	0.0596	0.0341

mechanisms with a Quick exponent ranging from 2 to 4 was necessary to obtain agreement between equal perceived contrast contours produced by their one dimensional difference of gaussian model and those obtained from contrast matching experiments. When response pooling was removed, implying that perceived contrast was mediated by the single mechanism producing the largest response, equal contrast contours produced by their model did not agree with those obtained experimentally. Our results agree in part with their conclusions. We reject the simple model corresponding to the removal of response pooling, but our two dimensional simulation appears to be much more selective in terms of Quick exponents that are acceptable. Our results indicate that Quick exponents of 4 produce perceived contrast responses that are too low while Quick exponents of 2 produce perceived contrast responses that are too high.

In summary, our results indicate that a model composed of simple non-linear mechanisms with constant bandwidth and gain across spatial frequencies can successfully predict perceived contrasts for a variety of two dimensional stimuli. When the mechanisms are Gabor sine functions, perceived contrast cannot be properly modelled by the response of a single mechanism (the simple model) but must be modelled, at least for the stimuli tested in this paper, by the pooled responses of several mechanisms tuned to different orientations and spatial frequencies.

The model discussed in this paper is a "first attempt" at two dimensional modelling, hence, the model is certainly not complete in its present form. Several further studies are anticipated. The model must be tested with other stimuli and the results compared to perceived contrast data derived from the same stimuli. Model performance for Gabor cosine mechanisms must be investigated since there is some indication that cortical mechanisms may exist in sine–cosine quadrature pairs (Pollen and Ronner, 1981). The effect of small changes in the power function exponent must be studied. Likewise, increased confidence in the generality of our conclusions regarding bandwidth and response pooling will require that a similar modelling effort be performed with mechanisms described by other mathematical functions such as difference of Gaussians.

REFERENCES

Cannon M. W. Jr (1979) Contrast sensation a linear function of stimulus contrast. *Vision Res.* **19**, 1045–1052.

Cannon M. W. Jr (1984) A study of stimulus range effects in free modulus magnitude estimation. *Vision Res.* **24**, 1049–1055.

Cannon M. W. Jr (1985) Perceived contrast in the fovea and periphery. *J. opt. Soc. Am.* **A2**, 1760–1768.

Daugman J. G. (1985) Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *J. opt. Soc. Am.* **A2**, 1160–1169.

De Valois R. L., Albrecht D. G. and Thorell L. G. (1982) Spatial frequency selectivity of cells in the macaque visual cortex. *Vision Res.* **22**, 545–559.

Drum B. (1986) Brightness estimation supports a two-channel model of achromatic brightness processing. *J. opt. Soc. Am.* **A3**, P84.

Georgeson M. A. and Sullivan G. D. (1975) Contrast constancy: deblurring in human vision by spatial frequency channels. *J. Physiol., Lond.* **252**, 627–656.

Gottesman J., Rubin G. and Legge G. E. (1981) A power law for perceived contrast in human vision. *Vision Res.* **21**, 791–799.

Kersten D. (1984) Spatial summation in visual noise. *Vision Res.* **12**, 1977–1990.

Klein S. A. and Levi D. M. (1985) Hyperacuity thresholds of 1 sec: theoretical predictions and empirical validation. *J. opt. Soc. Am.* **A2**, 1170–1190.

Legge G. E. and Foley J. M. (1980) Contrast masking in human vision. *J. opt. Soc. Am.* **70**, 1458–1471.

Parker A. and Hawken M. (1985) Capabilities of monkey cortical cells in spatial-resolution tasks. *J. opt. Soc. Am.* **A2**, 1101–1114.

Pollen D. A. and Ronner S. F. (1981) Phase relationships between adjacent simple cells in the visual cortex. *Science* **212**, 1409–1411.

Quick R. F. Jr (1974) A vector-magnitude model of contrast detection. *Kybernetic* **16**, 65–67.

Robson J. G. and Graham N. (1981) Probability summation and regional variation in contrast sensitivity across the visual field. *Vision Res.* **21**, 409–418.

Stevens S. S. (1956) The direct estimation of sensory magnitudes-loudness. *Am. J. Physiol.* **LXIX**, 1–25.

Swanson W. H., Wilson H. R. and Giese S. C. (1984) Contrast matching data predicted from contrast increment thresholds. *Vision Res.* **24**, 63–75.

Takahashi S. and Ejima Y. (1984) Dependence of apparent contrast for a sinusoidal grating on stimulus size. *J. opt. Soc. Am.* **A1**, 1197–1201.

Watson A. B. (1982) Summation of grating patches indicates many types of detector at one retinal location. *Vision Res.* **22**, 17–25.

Watson A. B. (1983) Detection and recognition of simple spatial forms. In *Physical and Biological Processing of Images* (Edited by Braddick O. J. and Sleigh A. C.), pp. 100–114. Springer, Berlin.

Webster M. A. and DeValois R. L. (1985) Relationship between spatial-frequency and orientation tuning of striate-cortex cells. *J. opt. Soc. Am.* **A2**, 1124–1132.

Wilson H. R. (1980) A transducer function for threshold and suprathreshold human vision. *Biol. Cybernet.* **38**, 171–178.

Wilson H. R. and Bergen J. R. (1979) A four mechanism model for threshold spatial vision. *Vision Res.* **19**, 19–32.

Wilson H. R., McFarlane D. K. and Phillips G. C. (1983) Spatial frequency tuning of orientation selective units estimated by oblique masking. *Vision Res.* **23**, 873–882.

Young R. A. (1985) The gaussian derivative model for

machine and biological image processing. Research publication GMR-5128, General Motors Research Laboratories, Warren, MI 48090, U.S.A.

APPENDIX 1

Nonlinear System Response and the Quick Summation Exponent

Note: The conclusions derived below are valid only when the system nonlinearity can be described as a simple power function.

The total system response, R, is the following function of the filter responses $S_{f,\theta}$;

$$R = \left[\sum_f \sum_\theta (|S_{f,\theta}|^p)^q\right]^{1/q},$$

where q is the Quick summation exponent and p is the power function exponent describing the non-linear transformation of the linear filter output. The summation symbol represents summation across spatial frequency and orientation as explained in the test.

Now change the contrast of the stimulus by a ratio c, so that each linear filter response is given by

$$S'_{f,\theta} = c\, S_{f,\theta}.$$

This produces a new system response

$$R' = \left[\sum_f \sum_\theta (|c\, S_{f,\theta}|^p)^q\right]^{1/q}$$

$$R' = \left[c^{pq} \sum_f \sum_\theta (|S_{f,\theta}|^p)^q\right]^{1/q} = c^p\, R.$$

Thus, the change in the system response is independent of the Quick summation exponent, q, and depends only on the perceived contrast exponent, p.

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

AD-A204 952

(2)

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) AAMRL-TR-88-033	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Armstrong Aerospace Medical Research Laboratory, AFSC, HSD 6b. OFFICE SYMBOL (If applicable) AAMRL/HEF	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State, and ZIP Code) Wright-Patterson AFB OH 45433-6573	7b. ADDRESS (City, State, and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
61102F	2313	V2	01

DTIC ELECTE MAR 0 2 1989 S D

11. TITLE (Include Security Classification)
Perceived Contrast and Stimulus Size: Experiment and Simulation (U)

12. PERSONAL AUTHOR(S)
Cannon, Mark W., Jr., Fullenkamp, Steven C.*

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Journal	FROM 10/85 TO 10/86	87 Dec 15	16

16. SUPPLEMENTARY NOTATION
*Systems Research Laboratories, Inc., 2600 Indian Ripple Road, Dayton OH 45440

17. COSATI CODES		
FIELD	GROUP	SUB-GROUP
05	08	

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
Contrast perception
Visual modeling
Spatial filtering

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Perceived contrast functions were determined for three different Gabor patch sizes using magnitude estimation and verified by contrast matching. While thresholds show a significant decrease with decreasing patch size, perceived contrasts are equal and independent of patch size for contrasts above 0.06. Contrast matching was also used to study the apparent contrast of two other spatially limited stimuli; the sum of two orthogonal 4 c/deg sine waves multiplied by a gaussian envelope and the sum of spatially adjacent positive and negative gaussians. Models of contrast perception, based on tuned Gabor spatial filters, were formulated and tested for agreement with our experimental data. A model that pools filter responses across spatial frequencies and orientations was found to be more in agreement with our data than a model that simply uses the response of a single, maximally excited, mechanism to mediate contrast perception. Optimum filter bandwidth was found to be about 1.1 octaves.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	UNCLASSIFIED

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code)	22c. OFFICE SYMBOL
Dr Mark W. Cannon, Jr.	(513) 255-8877	AAMRL/HEF

DD Form 1473, JUN 86 — Previous editions are obsolete. — SECURITY CLASSIFICATION OF THIS PAGE

UNCLASSIFIED

89 3 07 007